Q1:

GPUs’ ability to carry out fast floating-point arithmetic in parallel significantly reduces the time required to train CNNs. [1]

Sufficiently large labeled data sets are used to train the network by backpropagation

Or

CNN requires large data sets for it to extract generalized features. [1]

Q2:

For down sampling:

The longer the strides, the smaller the image after convolution will be. [1]

Pooling layers are used to reduce the sample size by only taking out part of the values [1]

For shift-invariant:

Multiple convolutional kernels ensure that the feature will appear at least at one place in the “convolutional volume” [1]

CNNs are trained such that the FC will be excited if the feature appears at any point in the convolution volume [1]

(diagrams are also accepted)

Q3

1. For finding a route of minimal length [1]

that can visit every station on the map [1]

1. It is not guaranteed to produce an optimal solution. [1]

Because it is a greedy algorithm that does not reconsider nodes. [1]